The Role of Context in the Performance of Iranian EFL Learners in Vocabulary Tests

Karim Sadeghi
English Language and Literature Department, Urmia University, Urmia, Iran
Email: k.sadeghi@urmia.ac.ir

Deniz Abdollahzadeh
English Language and Literature Department, Urmia University, Urmia, Iran
Email: Abdollahzadeh.mt@gmail.com

Abstract—This study investigated the role of context in the performance of Iranian EFL learners in vocabulary tests. The study involved both qualitative and quantitative methods of research. Four intact classes at intermediate level were selected. Participants were given the vocabulary section of an institutional TOEFL to make sure they are appropriately placed. Outliers were omitted and 60 students were selected. Four vocabulary tests each providing varying degrees of context were constructed and used to assess the participants’ lexical knowledge. These tests included word definition matching test with no context, limited context test with unconnected sentences, reduced context test with cohesive but not coherent sentences and a cloze test providing the whole context. Following the tests, students were given a questionnaire to write their comments on these tests. The application of One-way repeated measures ANOVA to quantitative data suggested that context affected the performance of test-takers in vocabulary tests with various contextual conditions. All the participants performed significantly better in reduced context test (sentential context condition) and had the weakest performance in the matching test (no context condition). Data coming from the questionnaire supported other findings: The majority of the participants agreed that reduced context test was the best for assessing their vocabulary knowledge. Further findings and implications are discussed in the paper.

Index Terms—vocabulary test, test context, Iranian EFL learners

I. INTRODUCTION

A test measures a person’s ability or knowledge. Every test requires a sample of performance. This performance can be used by the tester to gauge behavior and infer general ability (Brown, 2004). Various types of tests (e.g., multiple-choice tests, cloze tests, matching tests, gap-filling tests, etc.) are used to assess learners’ vocabulary knowledge in different languages. Each of the above mentioned tests provides various degrees of context for testees. Whether testees attend to contextual clues in taking vocabulary tests has been a matter of interest for many researchers such as Read (2000), Read and Chapelle (2001), Rivers (1981) and Schmitt (2001).

Communicating in natural communication situations, learners are faced with unfamiliar lexical items, for which they must find ways of guessing the meaning in order to comprehend the message adequately. Classroom teachers can solve the problem of unfamiliar lexical items by pre-teaching them, translating, providing synonym, dictionary definition or using peer assistance; however, in natural communication situations the most important strategy according to Read (2000) is inferring the meaning from information available in the text itself. This shows the importance of using context to guess the meanings of unfamiliar words, both in natural communication situations and in language vocabulary tests.

The history of language testing has undergone many changes. Language testing trends have followed the changes in language teaching methodology which have been influenced by theories in linguistics and psychology. Based on different views in language teaching, linguistics and psychology, four major trends are recognized in language testing. During the time when grammar translation method of teaching was fashionable, linguistics followed a traditional approach in which there was no concern for objectivity or reliability. At this time, tests comprised passage translation and essay writing which were scored subjectively. In 1950s, with the genesis of behaviorism in psychology and structuralism in linguistics, there was a major change in language tests. Language teachers adopted the audio-lingual method of teaching which in turn made use of discrete-point tests. In discrete-point tests, every language skill was tested separately in different items (Chun-mei, 2007). When the attention of linguists and psychologists shifted from structuralism and behaviorism to generativism and cognitivism, teachers adopted cognitive-code approach. As a result, there was a shift in language tests from discrete-point tests to integrative tests. Integrative tests concentrated on the communicative use of language and there was an attempt to create tasks similar to real life situation tasks. Unlike discrete-point tests, these tests assessed all sub-skills of language at the same time (Farhady, Ja’farpur & Birjandi, 1994). During the time when the focus of linguistics and psychologists was on psycho-linguistics and pragmatism, functional
approach in language testing emerged. Functional-communicative tests attempt to assess the testees’ performance in certain situations. For instance, testees are presented with integrated texts for which they have to interact with and produce integrated written or spoken texts (Birjandi, Bagerdoust and Mossalanejad, 2000).

As far as vocabulary testing is concerned, “whether to adopt context is a heat-debated topic” (Chun-mei, 2007, p. 42). Dixon-Krauss (2001) believes that as trends change in vocabulary teaching, changes in the ways learners are assessed in vocabulary tests appear as well. In the early years of objective testing it was assumed by the scholars that vocabulary tests must always present the target items in isolation. As a result of presenting target words in isolation, some problems occurred. A word can have different meanings in different contexts and can be used as more than one part of speech. When tested in isolation, it was impossible for testees to decide which meaning and which part of speech was intended by the test designer (Read, 2000). Sims (1929) attempted to solve the problem by including the words that only have one meaning in the test; however, this is not a feasible solution in normal situations of vocabulary tests. Miller (1999) asserts that the meanings of words are associated with different contexts. To elaborate the concept, he provides an example. In the sentence, That was a good shot, the word shot may have different meanings based on various contexts in which it is used. Shot can refer to a drink in a bar; it can refer to an image for a photographer and to an athlete; it can also refer to a hit, stroke, or kick of the ball. He reports that it is not possible for learners to gain such knowledge through a decontextualised learning task. “In the field of language teaching, where it has become commonplace to define both goals and methods primarily in communicative terms, it can seem rather suspect to be assessing vocabulary knowledge through decontextualised test items” (Read, 2007, p. 121).

In his experiment, Chun-mei (2007) investigated the relationship between the amount of context and performance in vocabulary tests. He attempted to find out the effect of zero context and sentence context on vocabulary testing. In addition, he investigated how different kinds of context in vocabulary tests affected subjects’ performance. Forty students from Zhejiang University participated in the study. The materials included two sets of vocabulary tests. The first test was composed of a 20-item test of 20 isolated single words with no context for which testees had to choose the most similar meaning from among the four alternatives provided. The second test (with 20 multiple-choice items) was in the form of a sentence which included the four alternatives provided. The testees were required to choose the most similar meaning for the underlined word in the sentences. The order of choices in each item and the order of the twenty items were rearranged in the second test in order to avoid the possibility of students’ memorizing answers in the first test and using them in the second test. All the sentences in test 2 were chosen from the examples of Collins COBUILD English Dictionary, Oxford Advanced Learner’s English-Chinese Dictionary, and Cambridge International Dictionary of English. Data analysis revealed a positive influence of sentence context on testees’ performance in vocabulary tests that context played an important role in helping testees to figure out the correct meanings of target words.

One comprehensive study on the role of contextual clues in test-takers’ performance in vocabulary testing was conducted by Uckun (2008). The study is perhaps the most widespread study of context contribution to EFL learners’ assessment of vocabulary gain. The researcher describes the development and trial of three measurement techniques each providing different degrees of context; word-definition matching test, gap-filling test and cloze test. The first test provides no context; the second one provides a reduced context; and the last one provides the whole context. The aim of the study was to investigate whether increasing the amount of context surrounding the target words from no context to reduced and whole context produced differences in the performance of testees for the same set of vocabulary items in differing contextual conditions. The participants of the study came from Gaziantep University in Turkey. One hundred eighty nine students from three language proficiency levels (elementary, intermediate and advanced levels) were given three types of tests at their own level of language proficiency in order to investigate how much context is needed for testees to be successful in vocabulary tests at different stages of linguistic ability. The researcher concluded that all subjects regardless of their language abilities benefited from the reduced context condition at sentence level. “The expectations of a greater success with the extended discoursal context at higher language ability levels was not supported by this study” (p. 21).

And finally, Webb (2008) investigated the effect of context on inferring the meanings of target words in vocabulary tests. Fifty Japanese university students encountered 10 target words in 3 sets of 10 short contexts that were rated on the amount of information available to infer the meanings of target words. Then another vocabulary test measured recall of form, recognition of form, recall of meaning and recognition of meaning. Based on the results, the group that read the texts containing more contextual clues had significantly higher scores on both tests of meaning which indicated that the quality of context rather than the number of encounters may have greater effect on in inferring the meaning of unknown words and gaining knowledge of meaning.

The importance of contextual clues in vocabulary tests for the prominent researchers mentioned above entails investigating the role of context in the performance of Iranian EFL learners in vocabulary tests. In order to check the efficiency of vocabulary teaching and learning, there is a need for the use of appropriately-constructed vocabulary tests. To serve such a purpose, this study aimed to investigate which type of vocabulary test (with various degrees of contextual clues) would be more appropriate at intermediate level. More specifically the current study sought to find answer to the following question:

1. Is there any relationship between the amount of context and the performance of Iranian EFL learners in vocabulary tests?
The following hypothesis was put forward for the above-mentioned question to be tested at 0.05 level of probability.

2. There is no relationship between the amount of context and the performance of Iranian EFL learners in vocabulary tests.

The current study adds to the previous ones in that apart from those contextual levels studied by Uckun (2008), a fourth contextual condition is added in which a text with superficial cohesive links but with no coherence is used. The present study also adds a new dimension to the study by giving a questionnaire at the end of the test to gather qualitative data on the subjects’ ideas about how much context is conducive to success in different tests which provide varying degrees of context.

II. MATERIALS AND METHODS

This study involved both qualitative and quantitative methods of research. Quantitative phase of research dealt with four measurement techniques or tests: a matching test, a limited context test, a reduced context test and a cloze test. Qualitative phase of research intended to seek attitudes of the test-takers on various types of tests they sat. Thus the study enjoyed a mixed-approach design.

The participants of this study consisted of 60 students within the age range of 15-20. At the time of study, all the participants were studying at an intermediate level in J.D.L.I. (Jahad-e-Daneshghahi Language Institute) in Urmia, Iran. The students are admitted to this institute using an in-house placement test which includes an interview section. They had different levels of general education (including those studying at schools and tertiary education centers) and had different purposes for learning English such as communicating abroad with natives or foreigners, continuing education or for entrance exam. At the time of research, the participants of the study were studying at Level 7 which is considered an intermediate level at this institute. Top Notch series (Slasow & Ascher 2006) are used as the main books for teaching English. The medium of instruction is English; however judicious use of Farsi is also allowed whenever needed.

Four vocabulary tests each providing varying degrees of context were constructed and used to assess the participants’ lexical knowledge. These tests included word definition matching test with no context, limited context test with unconnected sentences, reduced context test with cohesive but not coherent sentences and a cloze test which provides the whole context. In order to understand participants’ reaction to different test types, a questionnaire was also developed and used in this research. Vocabulary section of an institutional Test of English as a Foreign Language (TOEFL) was used to make sure that all participants are homogeneous in terms of lexical knowledge. The following gives detailed information on these instruments.

The matching test: The basic task for participants in this test was to match the target words with their dictionary definitions. No synonyms or antonyms were used in order to preserve the homogeneity of materials. Twelve words with approximately the same frequency levels were selected from Top Notch book (part 3B). These words were grouped in one column and definitions in another. All the frequencies were checked against Collins COBUILD Advanced Learner’s English Dictionary. Some of the definitions were provided within the book; for other definitions, Longman Dictionary of Contemporary English (1995) was used. To prevent the predictability of the last item, unequal number of items was used in each column; therefore, two extra definitions were added to the left column. The shorter column was placed to the right and the longer one to the left since “placing the shorter column in the right facilitates the testee’s job in that he reads the long column first and moves through the shorter column to find the corresponding items” (Farhady, et al., 1994, p. 92). Testees were expected to match the items in the left column with the definitions in the right column and write the corresponding letter of each definition in front of each word in the right column. In this study, matching test was considered the most context-reduced type of test. The prepared items were piloted in all tests with similar participants before they were given to the real participants. An example of the matching test is provided below.

Honeymoon….c…

- a. a safe place
- b. good with hand
- c. vacation taken by two newly weds

Limited context test: Limited context test is a type of gap-filling test that assesses the testees’ knowledge using a number of propositionally unlinked sentences, each with a blank. Testees were expected to choose the most appropriate choice from among the four choices provided, to fill in the blank in each item. All the alternatives and distracters in all tests were selected by the researcher from Top Notch book series and were checked against Collins COBUILD Advanced Learner’s English Dictionary to make sure that they had approximately the same frequency levels. Testees were expected to mark the correct choice in each item by circling it. An item used in limited context test is provided below as an example.

The government has been ………. for failing to decrease the crimes.
- a. suggested
- b. appeared
- c. handled
- d. criticized

Reduced context test: Unlike the limited context test, this test assesses the testees’ lexical knowledge with a number of propositionally connected sentences, each with a blank. Each sentence provides four alternatives for testees from which they are asked to choose the most appropriate one to fill each blank. Two examples of reduced context test are provided below to show how these sentences are propositionally connected with each other.

I. The steam locomotives ………. transportation of products for long hours.
The cloze test: The passage for the cloze test was also taken from the Top Notch books. In choosing the passage, the opinions of teachers at the institute (J.D.L.I) was also taken into account. Teachers had the choice of 10 reading passages from a Top Notch book and were asked to choose the one that could be used as the best achievement test for students at intermediate level. Teachers were also asked to select the words from the passage, which were supposed to be deleted from the passage and to be assessed through cloze procedure in multiple choice formats. After words were selected by the teachers, they were checked against Collins COBUILD Advanced Learner’s English dictionary to make sure that they all had approximately the same frequency levels. Teachers selected two passages to be used in the pilot test. After the pilot test, the one which was easier for the participants was selected to be used as the target cloze test passage for real participants of the study. The cloze test was considered the most context-embedded one in this research. Two examples of the cloze test items are provided below.

Every day we see and hear about terrorism and ……1….. acts committed against innocent people for religious or political ……2…..

1. a. corruptible    b. creative    c. innovative    d. violent
2. a. bribes    b. reasons    c. contracts    d. tickets

The questionnaire: A short questionnaire was prepared to elicit data on the testees’ opinions about four various types of vocabulary tests. It was aimed to find out which vocabulary test was more appropriate to assess the learners’ lexical knowledge at intermediate level and to find out what kind of contextual clues they actually attended to in their lexical decisions. An example of the questionnaire item is provided below.

In your opinion, which test assesses your lexical knowledge better?
a. Matching test    b. Limited context test
    c. Reduced context test    d. Cloze test

TOEFL (Test of English as a Foreign Language): The main purpose of this test is to assess whether foreign students planning to study in a tertiary institution where English is the medium of instruction have a sufficient level of language proficiency to study the academic courses without having language-related difficulties. The sole purpose of using this test in this study was to assure that all the participants were approximately at the same level of lexical proficiency; therefore, only vocabulary section of this test was used in this study.

In order to conduct this research, four intact classes, two male and two female classes at intermediate level were selected from Jahad-e-Daneshgah Language Institute. Although all the participants had already taken a placement test on arrival to the institute and progressed to Level 7 based on final exams, they were further given vocabulary section of an institutional TOEFL to make sure that all participants were homogeneous in terms of lexical knowledge and were appropriately placed at their own proficiency levels. Based on the means of the TOEFL, outliers were omitted and 60 testees (out of a total of 78) were selected to take part in the study. First they were given a cover letter to assure them that the personal information they provide would not be disclosed to anyone else, and all the information either in the questionnaire or the tests would be used for the purpose of this research only. Then each of the participants was given four different vocabulary tests (as described above), each with a different contextual condition. Each test consisted of 12 items, each having one score; thus, the maximum score for each test was 12 and the total score for all the tests was 48. Testees were not penalized for their wrong answers. The allocated time to answer all the questions was one hour. The amount of time was determined based on a pilot test which was carried out with similar participants. After the tests, each participant was given the questionnaire in order for the participants to give their comments on four various kinds of vocabulary tests they had just taken and in order to provide an opportunity for the researchers to find out what kind of contextual clues they actually attended to in their lexical decisions.

For analyzing the data quantitatively, One-way repeated measures ANOVA was used to examine the relationship between the role of context and the performance of all testees in vocabulary tests. Data was analyzed using Statistical Package for Social Sciences (SPSS). The analysis of the qualitative data was carried out through description and interpretation techniques.

III. Results

The present study investigated the relationship between the amount of context and the performance of Iranian EFL learners in vocabulary tests. It was hypothesized that there is no relationship between the amount of context and the performance of testees in vocabulary tests. Table 1 shows descriptive statistics for all test-takers in four contextual conditions (matching test, limited context test, reduced context test and cloze test).
In order to find out whether these means are statistically significant or not and to test the hypothesis mentioned above, test-takers’ scores in vocabulary tests with various contextual conditions were analyzed through one-way repeated measures ANOVA.

### TABLE 1

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limited context test</td>
<td>7.2759</td>
<td>2.726695</td>
<td>60</td>
</tr>
<tr>
<td>Matching test</td>
<td>6.0517</td>
<td>3.199835</td>
<td>60</td>
</tr>
<tr>
<td>Cloze test</td>
<td>6.8621</td>
<td>3.116780</td>
<td>60</td>
</tr>
<tr>
<td>Reduced context test</td>
<td>8.2759</td>
<td>2.514605</td>
<td>60</td>
</tr>
</tbody>
</table>

Table 1 shows the means of test takers in vocabulary tests with various contextual conditions.

### TABLE 2

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>context</td>
<td>.404</td>
<td>12.888*</td>
<td>3.000</td>
<td>57.000</td>
<td>.000</td>
<td>.404</td>
</tr>
<tr>
<td>Pillai’s Trace</td>
<td>.596</td>
<td>12.888*</td>
<td>3.000</td>
<td>57.000</td>
<td>.000</td>
<td>.404</td>
</tr>
<tr>
<td>Wilks’ Lambda</td>
<td>.678</td>
<td>12.888*</td>
<td>3.000</td>
<td>57.000</td>
<td>.000</td>
<td>.404</td>
</tr>
<tr>
<td>Hotelling’s Trace</td>
<td>.678</td>
<td>12.888*</td>
<td>3.000</td>
<td>57.000</td>
<td>.000</td>
<td>.404</td>
</tr>
<tr>
<td>Roy's Largest Root</td>
<td>.678</td>
<td>12.888*</td>
<td>3.000</td>
<td>57.000</td>
<td>.000</td>
<td>.404</td>
</tr>
</tbody>
</table>

According to Table 2, the relationship between the test-takers’ performance and the amount of context in vocabulary tests was statistically significant. There was a significant effect for context, Wilk’s Lambda = 0.596, F (3, 57) = 12.888, p = .000, partial eta squared = .404, which is a very large effect size.

Pairwise comparisons indicated that all test-takers performed significantly better in reduced context test (sentential level condition) and had the weakest performance in the matching test (no context condition). Based on the results of the study, it can be concluded that context affects the performance of all testees in vocabulary tests with various contextual conditions; therefore, the hypothesis that there is no relationship between the amount of context and the performance of testees in vocabulary tests was rejected.

The analysis of qualitative data through description and interpretation techniques revealed that most of the participants of the study preferred reduced context test as the best way of assessing their vocabulary knowledge. When they were asked which of the vocabulary tests (matching test, limited context test, reduced context test and cloze test) assessed their lexical knowledge better, they all believed that multiple-choice gap-filling tests (limited and reduced context tests) were the best way of assessing their knowledge of language since these kinds of tests were similar to the university entrance exam tests, providing a chance for them to practice for university entrance exams. It seems obvious that the washback candidates expected has affected their attitudes toward vocabulary testing.

The findings of qualitative analysis was in line with quantitative findings indicating that lack of context in matching test hindered their chance of inferring unknown words; on the other hand, when the context was too long (i.e., cloze test) they got confused. The findings support those by Read (2000) who believes that the length of context does not necessarily make it easier for the students to understand the meaning of unknown words; it is the level of language proficiency that determines students’ ability to use the contextual clues; hence, partial knowledge of learners in developing context will lead to wrong guesses. Therefore, in this research the idea that as the context becomes larger, students can guess the meaning of unknown lexis better was not substantiated.

### IV. DISCUSSION

Generally speaking, in recent years more and more tests have begun to test vocabulary in context. For instance, while in the past TOFEL vocabulary tests provided only a word to be tested and four choices of single words, sentential context was later adopted in TOFEL vocabulary tests. Later, from 1995, TOFEL began to test vocabulary in reading materials. Other standard tests, such as Common English Test (CET), Preliminary English test (PET), Graduate Management Admission Test (GMAT) and International English Language Testing System (IELTS) also began adopting context at sentential or paragraph levels (Chun-mei, 2007).

Regarding the research question, whether there is a relationship between the amount of context and the performance of Iranian EFL learners in vocabulary tests, the statistically significant correlation (p=0.000), indicated that the type of contextual condition in vocabulary tests can affect the performance of test-takers at intermediate level of language proficiency. The findings showed that contextual condition should be considered one of the major variables in designing vocabulary tests since it can affect students’ performance; nonetheless, the amount of context must be determined by the test-takers’ level of language proficiency since it was concluded in this study that learners at intermediate level can...
only make use of context at sentential level which was presented through two types of gap-filling tests (reduced and limited context tests). The results also indicated that when sentences were related to each other cohesively (reduced context test) the possibility of inferring the meaning of unknown words for test-takers increased. The cloze test with the highest amount of context and matching test with no context did not result in higher scores for any of the test-takers.

The findings of the present study are in line with those reported by Chun-mei (2007), and Uckun (2008) in that sentence context leads to better performance among testees in inferring the meaning of unknown words in vocabulary tests. On the other hand, the findings by Webb (2008) appear to contradict those in this study. He concluded the larger the context and the more the contextual clues, the better the performance of test-takers in vocabulary tests. In his study, the group that read the texts containing more contextual clues had significantly higher scores on both tests of meaning.

V. CONCLUSION

Differences in the performance of testees in different vocabulary tests e.g., matching test, limited context test, reduced context test and cloze test could be due to varying amount of contextual conditions provided in these tests. Consequently, in order to help test-takers feel at ease and demonstrate their full lexical knowledge in vocabulary tests, teachers had better try to construct the most suitable vocabulary tests with the right amount of contextual conditions for learners at different proficiency levels. The findings showed that contextual condition should be considered one of the major variables in designing vocabulary tests since it can affect students’ performance; nonetheless, the amount of context must be determined by the test-takers’ level of language proficiency since it was observed that learners at intermediate level can only make use of context at sentential level which was presented through two types of gap-filling tests (reduced and limited context tests). The results also indicated that when sentences were related to each other cohesively (reduced context test), the possibility of inferring the meaning of unknown words for test-takers increased.

A limitation of the study is that since the participants at intermediate level came from different classes of the same institute and consequently had to take the tests under different conditions, environmental extraneous variables might have affected their performance. Among the environmental variables that could possibly contribute to differential performance of the participants in this study are “noise, temperature, adequacy of light, time of day, and seating arrangements” (Brown, 1995, p. 30). Another limitation of this study is related to the fact that because of limited number of students in higher levels, the study was conducted only with students at intermediate level not advanced ones. Replication of this study with participants at other levels of language proficiency such as elementary and advanced would support or challenge our findings. The study can also be replicated with participants with different age ranges. Measures other than the ones used in this study can be investigated for their contextual contribution. Investigating the role of contextual conditions in different language skills such as grammar will further contribute to the field and improve our understanding of testing issues, making the results generalizable to more contexts and more skill areas.

REFERENCES


Karim Sadeghi has a PhD from the University of East Anglia, UK, and is an assistant professor of TESOL at Urmia University, Iran. His publications have appeared in Asian EFL Journal, The Asia Pacific Education Researcher, The Modern Language Journal, Canadian Journal of Linguistics and TESL Canada Journal among others.
Deniz Abdollahzadeh is an MA candidate in TEFL at Urmia University, Iran. She received her B.A. in English Language and Literature from Urmia University. Her main research interests are language testing and assessment. She is currently teaching English at the Language Center of Urmia University and Jahad-e-Daneshgahi Language Institute.